**A Project report on**

#### PNEUMONIA DETECTION USING DEEP LEARNING

A Dissertation submitted in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**In**

**Computer Science and Engineering**

Submitted by

**P. Tanuja Reddy**

**(20H51A05A3)**

**N. Aditya**

**(20H51A05D1 T.Ruchitha**

**(20H51A05M1)**

Under the esteemed guidance of

**Ms.M.Srikala**

Assistant professor



### Department of Computer Science and Engineering

**CMR COLLEGE OF ENGINEERING AND TECHNOLOGY**

(An Autonomous Institution under UGC & JNTUH , Approved by AICTE, Permanently Affiliated to JNTUH, Accredited by NBA.)

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

###### 2020- 2024

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



## CERTIFICATE

This is to certify that the Project report entitled **"PNEUMONIA DETECTION USING DEEP LEARNING"** being submitted by **P.Tanuja Reddy(20H51A05A3), N.Aditya(20H51A05D1) ,T.Ruchitha(20H51A05M1)** in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

###### Ms. M. Srikala Dr. Siva Skandha Sanagala

**Assistant Professor Associate Professor and HOD**

**Dept. of CSE Dept. of CSE**

## Acknowledgement

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to**Ms.M.Srikala, Assistant Professor** , Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

**P.Tanuja Reddy (20H51A05A3)**

**N.Aditya (20H51A05D1)**

**T.Ruchitha (20H51A05M1)**

### TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| **NO.** | **TITLE** | **PAGE NO.** |
|  | **LIST OF FIGURES**  **LIST OF ABBREVIATIONS** | i i |
|  | **ABSTRACT** | 3 |
| **1** | **INTRODUCTION** |  |
|  | 1.1 Motivation | 4 |
|  | 1.2 Objective | 5 |
| **2** | **BACKGROUND WORK** |  |
|  | 2.1 Domain Introduction | 6 |
|  | 2.2 Deep Learning | 6 |
|  | 2.3 Convolution Neural Network | 8 |
|  | 2.4 Transfer Learning Models | 9 |
| **3** | **EXISTING SYSTEM** |  |
|  | 3.1 VGG | 10 |
|  | 3.1.1 code and output | 11-13 |
|  | 3.2 Inception V3 | 14 |
|  | 3.2.1 code and output | 15 |
|  | 3.3 Densenet | 16 |
|  | 3.3.1 code and output | 17 |
| **4** | **PROPOSED SYSTEM** | 24 |
|  | 4.1 CNN segmentation+connected components | 25 |
|  | 4.2 Diagrams | 27 |
|  | 4.3 Code and Output | 28 |
| **5** | **DESIGNING** | 30 |
| **6** | **RESULTS AND DISCUSSION** | 33 |
| **7** | **CONCLUSION AND FUTUREWORK** | 40 |
| **8** | **REFERENCES** | 46 |

**List Of Figures**

|  |  |  |
| --- | --- | --- |
| **FIGURES** | **TITLE** | **PAGE NO.** |
| 1.1 | Chest x-rays images | 5 |
| 2.2 | Bayes Classifier | 7 |
| 2.3 | Convolution neural network | 9 |
| 3.1 | VGG | 13 |
| 3.2 | Inception v3 | 16 |
| 3.3 | Densenet 121 | 18 |
| 4.1 | CNN segmentation | 22 |
| 5.1 | Designing | 31 |
| 5.2 | Cnn components | 32 |
| 6.1 | IOU Loss | 34 |
| 6.2 | Result | 35 |

#### LIST OF ABBREVIATIONS

|  |  |  |
| --- | --- | --- |
| **S.no** | **ABBREVIATIONS** | **EXPANSION** |
| 1. | AUC | Area Under Curve |
| 2. | CXR | Chest X-Ray |
| 3. | CNN | Convolution Neural Network |
| 4. | WHO | World Health Organization |
| 5. | RELU | Rectified Linear Unit |
| 6. | VGG | Visual Geometric Group |
| 7. | DNN | Deep Neural Network |
| 8. | DENSENET | Dewsely Connected Neural Network |
| 9. | YOLO | You Only Look Once |
| 10. | RCNN | Region Based Convolution Neural Network |
| 11. | IOU | Intersection over Union |

**ABSTRACT**

Pneumonia causes the death of around 700,000 children every year and affects 7% of the global population. Chest X-rays are primarily used for the diagnosis of this disease. However, even for a trained radiologist, it is a challenging task to examine chest X-rays. There is a need to improve the diagnosis accuracy .In this work, an efficient model for the detection of pneumonia trained on digital chest X-ray images is proposed, which could aid the radiologists in their decision making process. The proposed weighted classifier is able to outperform all the individual models. Finally, the model is evaluated,not only in terms of test accuracy, but also in the AUC score.

Early diagnosis of pneumonia is crucial to cure the disease completely. Examination of X-ray scans is the most common means of diagnosis, but it depends on the interpretative ability of the radiologist and frequently is not agreed upon by the radiologists. The main objective is to predict and develop an deep learning to diagnose pneumonia in images and differentiate it from pneumonia and non pneumonia diseases. The other objective is to achieve higher accuracy for prediction. We develop an algorithm which detects pneumonia from frontal-view chest X-ray images at a level exceeding practicing radiologist. We also show that an easy extension of our algorithm to detect multiple diseases outperforms previous state of the art on ChestX-ray14, the most important publicly available chest X-ray dataset.

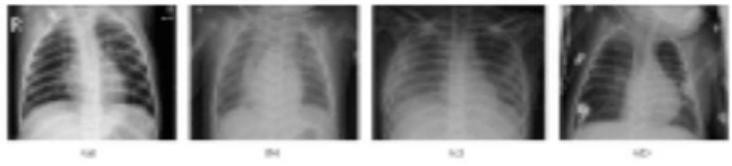
1

# CHAPTER1 INTRODUCTION

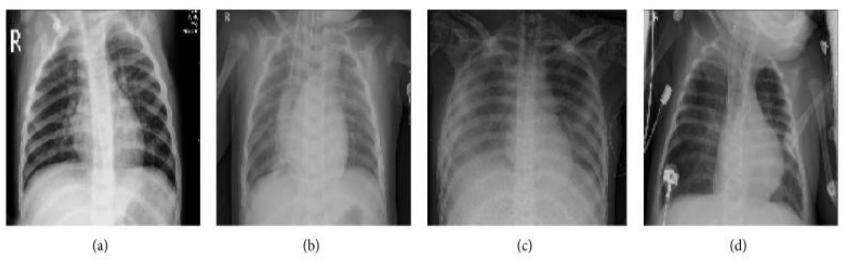
#### INTRODUCTION

The risk of pneumonia is immense for many, especially in developing nations where billionsface energy poverty and rely on polluting forms of energy. The WHO estimates that over 4 million premature deaths occur annually from household air pollution-related diseases including pneumonia. Over 150 million people get infected with pneumonia on an annual basis especially children under 5 years old.

While common, accurately diagnosing pneumonia is a tall order. It requires review of a chest radiograph (CXR) by highly trained specialists and confirmation through clinical history, vital signs and laboratory exams. Pneumonia usually manifests as an area or areas ofincreased opacity on CXR. However, the diagnosis of pneumonia on CXR is complicated because of a number of other conditions in the lungs such as fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, or post-radiation or surgical changes.To overcome this problem, a novel but simple model is introduced to automatically perform optimal classification tasks with deep neural network architecture. In recent times, CNN-motivated deep learning algorithms have become the standard choice for medical image classifications although the state- of-the-art CNN-based classification techniques pose similar fixated network architectures of the trial-and-error system which have been their designing principle. This study proposes a conceptually simple yet efficient network model to handle the pneumonia classification problem as shown in Figures 1 and 2.



**FIG-1: X-rays of non pneumonia**



**FIG-2:X- rays of pneumonia**

## Motivation:

As previously mentioned, pneumonia affects a large number of individuals, especially children, mostly in developing and underdeveloped countries characterized by risk factors such as overcrowding, poor hygienic conditions, and malnutrition.

Early diagnosis of pneumonia is crucial to cure the disease completely. Examination of X-ray scans is the most common means of diagnosis, but it depends on the interpretative ability of the radiologist and frequently is not agreed upon by the radiologists.

## Objective:

The main objective is to predict and develop an deep learning to diagnose pneumonia in images and differentiate it from pneumonia and non pneumonia diseases. The other objective is to achieve higher accuracy for prediction.

# CHAPTER 2 BACKGROUND WORK

#### DOMAIN INTRODUCTION:

The domain under which the study has been done are Deep Learning Models and a widely used library for Deep Learning Models known as Keras. More about this domain is explained in the below chapters.

#### DEEP LEARNING :

Deep learning is based on the branch of machine learning, which is a subset of artificial intelligence. Since neural networks imitate the human brain and so deep learning will do. In deep learning, nothing is programmed explicitly. Basically, it is a machine learning class that makes use of numerous nonlinear processing units so as to perform feature extraction as well as transformation. The output from each preceding layer is taken as input by each one of the successive layers.

Deep learning models are capable enough to focus on the accurate features themselves by requiring a little guidance from the programmer and are very helpful in solving out the problem of dimensionality. [Deep learning algorithms](https://www.javatpoint.com/deep-learning-algorithms) [a](https://www.javatpoint.com/deep-learning-algorithms)re used, especially when we have a huge no of inputs and outputs. Since deep learning has been evolved by the [machine learning](https://www.javatpoint.com/machine-learning) ,which itself is a subset of artificial intelligence and as the idea behind the [artificial intelligence](https://www.javatpoint.com/artificial-intelligence-tutorial) is to mimic the human behavior, so same is "the idea of deep learning to build such algorithm that can mimic the brain".

Deep learning is implemented with the help of Neural Networks, and the idea behind the motivation of [Neural Network](https://www.javatpoint.com/artificial-neural-network) is the biological neurons, which is nothing but a brain cell

“Deep learning is a collection of statistical techniques of machine learning for learning feature hierarchies that are actually based on artificial neural networks.”

Many of the early successes of AI took place in relatively sterile and formal environments and

did not require computers to have much knowledge about the world. For example, IBM’s Deep Blue chess- playing system defeated world champion Garry Kasparov in 1997 (Hsu, 2002). Chess is of course a very simple world, containing only sixty-four locations and thirty-two pieces that can move in only rigidly circumscribed ways. Devising a successful chess strategy is a tremendous accomplishment, but the challenge is not due to the difficulty of describing the set of chess pieces and allowable moves to the computer. Chess can be completely described by a very brief list of completely formal rules, easily provided ahead of time by the programmer.

Ironically, abstract and formal tasks that are among the most difficult mental undertakings for a human being are among the easiest for a computer. Computers have long been able to defeat even the best human chess player but only recently have begun matching some of the abilities of average human beings to recognize objects or speech. A person’s everyday life requires an immense amount of knowledge about the world. Much of this knowledge is subjective and intuitive, and therefore difficult to articulate in a formal way. Computers need to capture this same knowledge in order to behave in an intelligent way.

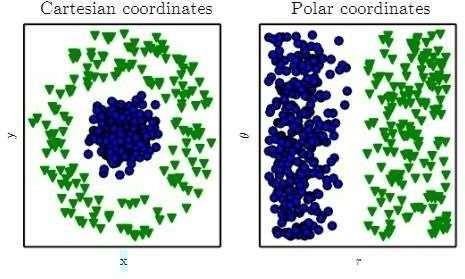
One of the key challenges in artificial intelligence is how to get this informal knowledge into a computer.

Several artificial intelligence projects have sought to hard-code knowledge about the world in formal languages. A computer can reason automatically about statements in these formal languages using logical inference rules. This is known as the knowledge base approach to artificial intelligence. None of these projects has led to a major success. One of the most famous such projects is Cyc (Lenat and Guha, 1989). Cyc is an inference engine and a database of statements in a language called CycL. These statements are entered by a staff of human supervisors. It is an unwieldy process. People struggle to devise formal rules with enough complexity to accurately describe the world. For example, Cyc failed to understand a story about a person named Fred shaving in the morning (Linde, 1992).

Its inference engine detected an inconsistency in the story: it knew that people do not have electrical parts, but because Fred was holding an electric razor, it believed the entity “Fred While Shaving” contained electrical parts. It therefore asked whether Fred was still a person while he was shaving. Predicting Lung Damage The difficult faced by systems relying on hard-coded knowledge suggest that AI systems need the ability to acquire their own knowledge, by extracting patterns from raw data. This capability is known

as machine learning.

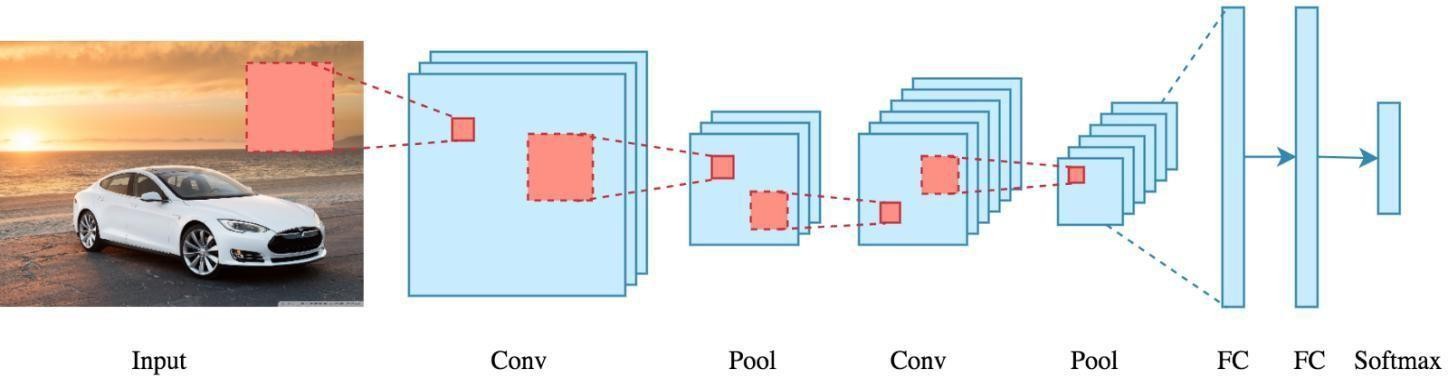
The introduction of machine learning enabled computers to tackle problems involving knowledge of the real world and make decisions that appear subjective. A simple machine learning algorithm called logistic regression can determine whether to recommend cesarean delivery (MorYosef et al., 1990). A simple machine learning algorithm called naive Bayes can separate legitimate email from spam e- mail.



**Fig. 3 Bayes Classifier**

#### CONVOLUTIONAL NEURAL NETWORK :

In deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.



**FIG-2: cnn architecture**

Images are made up of pixels. Each pixel is represented by a number between 0 and 255. Therefore each image has a digital representation which is how computers can work with images.

There are 4 major operations in CNN image detection/classification.

* Convolution
* Max pooling
* Flattening
* Fully connected layer

###### Convolution

The main purpose of a convolutional layer is to detect features or visual features in images such as edges, lines, color drops, etc. This is a very interesting property because, once it has learned a characteristic at a specific point in the image, it can recognize it later in any part of it.CNN’s make use of filters (also known as kernels, feature detectors), to detect features, such as edges, are present throughout an image. A filter is just a matrix of values, called weights that are trained to detect specific features.

###### Padding

One tricky issue when applying convolutional layers is that we tend to lose pixels on the perimeter ofour image. Since we typically use small kernels, for any given convolution, we might only lose a fewpixels, but this can add up as we apply many successive convolutional layers.

###### Max Pooling

After ReLU comes to a pooling step, in which the CNN downsamples the convolved feature (to save on processing time), while also reducing the size of the image. This helps reduce overfitting, which would occur if CNN is given too much information, especially if that information is not relevant in classifying the image.

###### Flattening

After multiple convolution layers and downsampling operations, the 3D representation of the image is converted into a feature vector that is passed into a multi-layer perceptron to output probabilities. The rows are concatenated to form a long feature vector. If multiple input layers are present, its rows are also concatenated to form an even longer feature vector.

###### Fully Connected Layer

In this step, the flattened feature map is passed through a neural network. This step is made up of the input layer, the fully connected layer, and the output layer. The fully connected layer is similar to the hidden layer in ANNs but in this case, it’s fully connected. The output layer is where we get the predicted classes. The information is passed through the network and the error of prediction is calculated. The error is then back propagated through the system to improve the prediction.

##### Transfer Learning Models :

From a deep learning perspective, the image classification problem can be solved through transfer learning. Transfer learning is a popular method in computer vision because it allows us to build accurate models in a timesaving way (Rawat & Wang 2017). With transfer learning, instead of starting the learning process from scratch, you start from patterns that have been learned when solving a different problem. This way you leverage previous learnings and avoid starting from scratch. Take it as the deep learning version of Chartres’ expression ‘standing on the shoulder of giants’. In computer vision, transfer learning is usually expressed through the use of pre-trained models. A pre-trained model is a model that was trained on a large benchmark dataset to solve a problem similar to the one that we want to solve. Accordingly, due to the computational cost of training such models, it is common practice to import and use models from published literature (e.g. VGG, Inception, MobileNet). When you’re repurposing a pre-trained model for your own needs, you start by removing the original classifier, then you add a new classifier that fits your purposes, and finally you have to fine-tune your model according to one of three strategies:

###### Train the entire model:

In this case, you use the architecture of the pre-trained model and train it according to your dataset. You’re learning the model from scratch, so you’ll need a large dataset.

###### Train some layers and leave the others frozen:

As you remember, lower layers refer to general features (problem independent), while higher layers refer to specific features (problem dependent). Here, we play with that dichotomy by choosing how much we want to adjust the weights of the network (a frozen layer does not change during training). Usually, if you’ve a small dataset and a large number of parameters, you’ll leave more layers frozen to avoid overfitting. By contrast, if the dataset is large and the number of parameters is small, you can improve your model by training more layers to the new task since overfitting is not an issue.

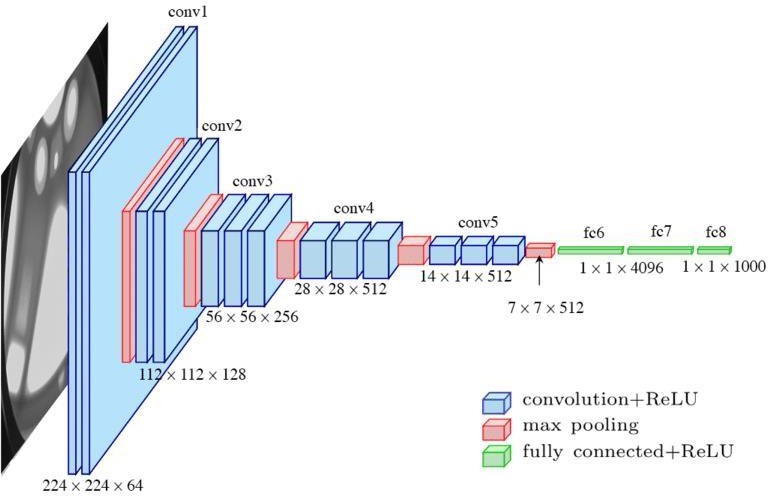
###### Freeze the convolutional base:

This case corresponds to an extreme situation of the train/freeze trade-off. The main idea is to keep the convolutional base in its original form and then use its outputs to feed the classifier. You’re using the pretrained model as a fixed feature extraction mechanism, which can be useful if you’re short on computational power, your dataset is small, and/or pre-trained model solves a problem very similar to the one you want to solve.

# CHAPTER 3 EXISTING SYSTEM

#### VGG :

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is an annual computer vision competition.Each year, teams compete on two tasks. The first is to detect objects within an image coming from 200 classes, which is called object localization. The second is to classify images, each labeled with one of 1000 categories, which is called image classification. VGG 16 was proposed by Karen Simonyan and Andrew Zisserman of the Visual Geometry Group Lab of Oxford University in 2014. The ImageNet dataset contains images of fixed size of 224\*224 and have RGB channels. So, we have a tensor of (224, 224, 3) . as our input. This model process the input image and outputs the a vector of 1000 values.



**FIG-3: Convolution Layers in CNN**

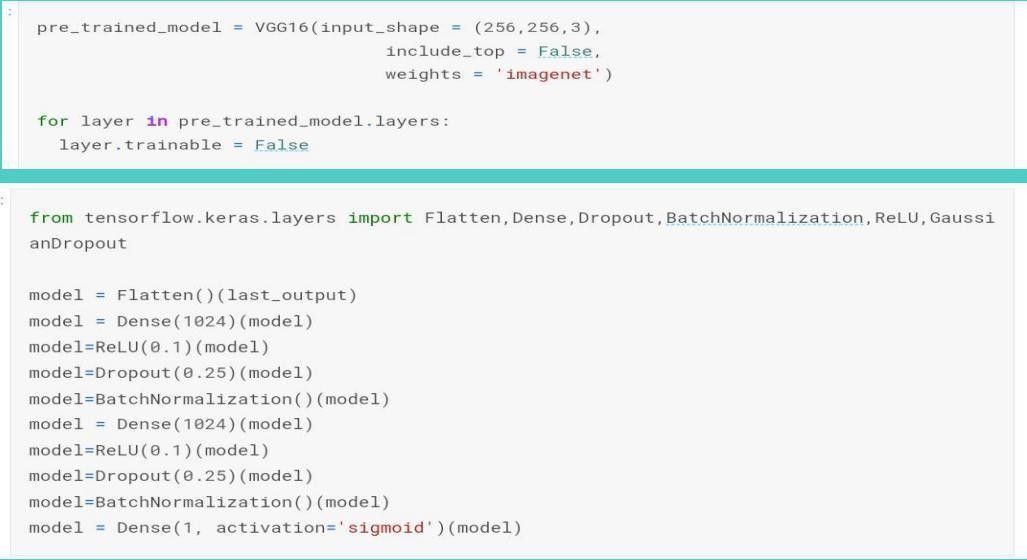
The input to the network is image of dimensions *(224, 224, 3)*. The first two layers have *64* channels of *3\*3* filter size and same padding. Then after a max pool layer of stride *(2, 2)*, two layers which have convolution layers of 256 filter size and filter size *(3, 3)*. This

followed by a max pooling layer of stride *(2, 2)* which is same as previous layer. Then there are *2* convolution layers of filter size *(3, 3)* and *256* filter. After that there are *2* sets of *3* convolution layer and a max pool layer. Each have *512* filters of *(3, 3)* size with same padding.This image is then passed to the stack of two convolution layers. In these convolution and max pooling layers, the filters we use is of the size *3\*3* instead of *11\*11* in AlexNet and *7\*7* in ZF-Net. In some of the layers, it also uses *1\*1* pixel which is used to manipulate the number of input channels. Thereis a padding of *1-pixel* (same padding) doneafter each convolution layer to prevent the spatial feature of the image.

After the stack of convolution and max-pooling layer, we got a *(7, 7, 512)* feature map. We flatten this output to make it a *(1, 25088)* feature vector.After this there are *3 fully* connectedlayer, the first layer takes input from the last feature vector and outputs a *(1, 4096)* vector, second layer also outputs a vector of size *(1, 4096)* but the third layer output a *1000* channels for *1000* classes of ILSVRC challenge, then after the output of 3rd fully connected layer is passed to softmax layer in order to normalize the classification vector. After the output of classification vector top-5 categories for evaluation. All the hidden layers use ReLU as its activation function. ReLU is more computationally efficient because it results infaster learning and it also decreases the likelihood of vanishing gradient problem.

VGG-16 was one of the best performing architecture in ILSVRC challenge 2014.It was the runner up in classification task with top-5 classification error of *7.32%* (only behind GoogLeNet with classification error *6.66%*). It was also the winner of localization task with *25.32%* localization error. It is very slow to train (the original VGG model was trained on Nvidia Titan GPU for 2-3 weeks).The size of VGG-16 trained imageNet weights is 528 MB.So, it takes quite a lot of disk space and bandwidth that makes it inefficient.

* + 1. **CODE:**



**OUTPUT:**

Epoch 16/20

/100 [==============================] - 374s 4s/step - loss: 0.9719 - accuracy:

0.5936 - val\_loss: 0.7614 - val\_accuracy: 0.7238 - lr: 1.2500e-04Epoch 17/2

/100/100 [==============================] - ETA: 0s - loss: 0.9606 - accuracy: 0.5897

Epoch 00017: ReduceLROnPlateau reducing learning rate to 6.25000029685907e-05.100/100 [==============================] - 372s 4s/step - loss: 0.9606 -

accuracy: 0.5897 - val\_loss: 1.5853 - val\_accuracy: 0.4509 - lr: 1.2500e-04

Epoch 18/20100/100 [==============================] - 409s 4s/step - loss:

0.9665 - accuracy: 0.5733 - val\_loss: 0.8606 - val\_accuracy: 0.7315 - lr: 6.2500e-05

Epoch 20/20100/100 [==============================] - 408s 4s/step - loss:

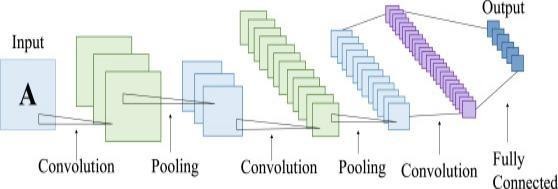
0.9568 - accuracy: 0.5804 - val\_loss: 1.0759 - val\_accuracy: 0.6492 - lr: 3.1250e-05

#### INCEPTION V3:

Inception models are a type of deep neural network (DNN) architecture developed by a researcher named Szegedy et al. for the first time in 2014, and the model was named as inception model . The structures of inception models and the conventional CNN model are different from each other in such a way that inception models are inception blocks which means lapping the same input tensor with multiple filters and concatenating their results. There are various versions of the inception models. In 2015, Szegedy et al. proposed a new version of the inception models named Inception- V3,which is an improved version of the previous versions of inception models, i.e., Inception-V1 and Inception-V2, and possesses more parameters.

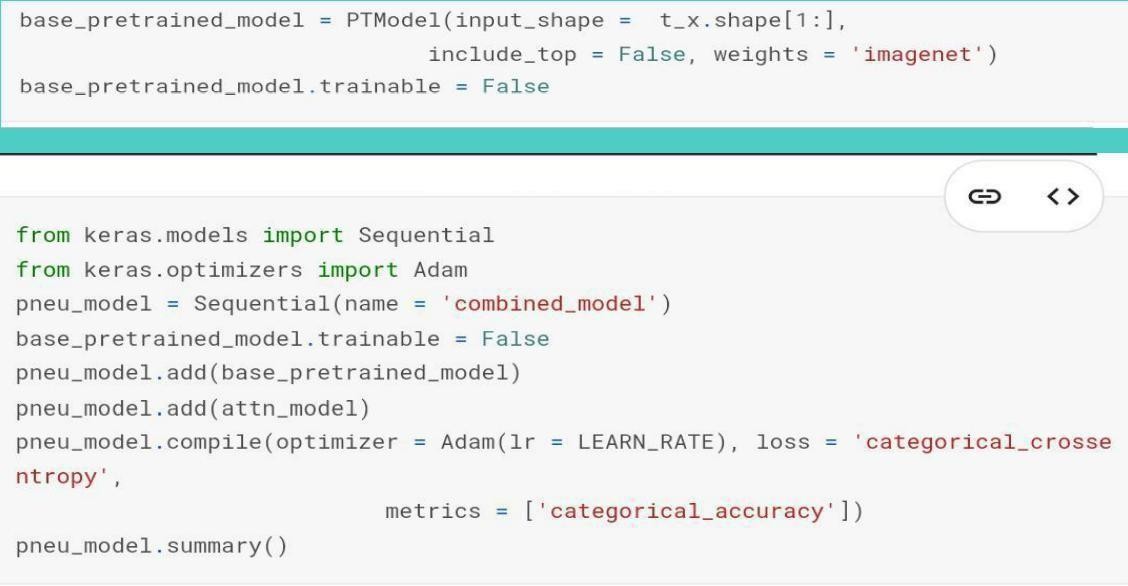
Inception-V3 contains a total of 24M parameters. The advancement in Inception-V3 was as follows:

(a) it factorizes the “n × n” convolution into asymmetric convolutions, i.e., 1 × n and n × 1, (b) it factorizes the 5 × 5 convolutions into two 3 × 3 convolutions, and (c) it replaces 7 × 7 convolutions toa series of 3 × 3 convolutions. Actually, it consists of a block of convolutional outputs are concatenated and sent to the next inception module.



**FIG-4: Inception architecture**

* + 1. **CODE:**



**OUTPUT:**

Epoch 15/20

250/250 [==============================] - 199s 797ms/step - loss: 0.9156 -

categorical\_accuracy: 0.5800 - val\_loss: 1.0927 - val\_categorical\_accuracy: 0.4767 Epoch 00015: val\_loss did not improve

Epoch 16/20 250/250 [==============================] - 185s 740ms/step - loss:

0.9074 - categorical\_accuracy: 0.5738 - val\_loss: 1.1385 - val\_categorical\_accuracy: 0.4733 Epoch 00016: val\_loss did not improve

Epoch 17/20 250/250 [==============================] - 155s 621ms/step - loss:

0.8966 - categorical\_accuracy: 0.5817 - val\_loss: 1.0366 - val\_categorical\_accuracy: 0.4967 Epoch 00017: val\_loss did not improve

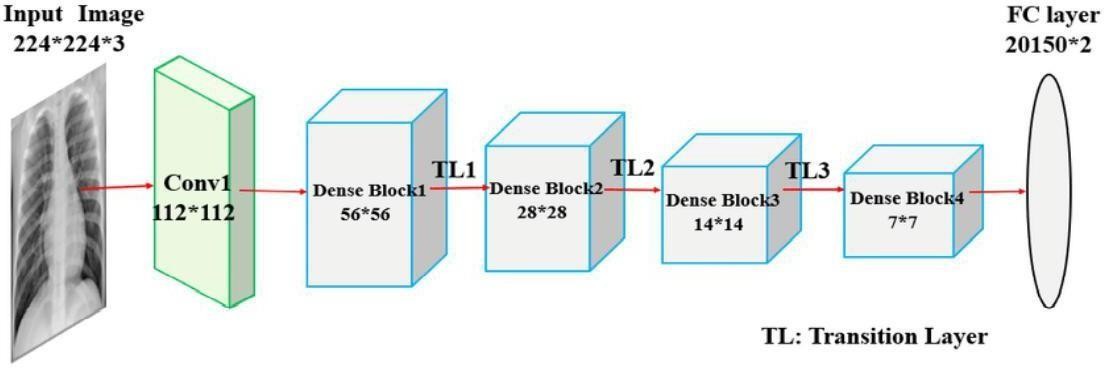
#### DENSENET 121

Densely Connected Convolutional Networks, DenseNets, are the next step on the way to keep increasing the depth of deep convolutional networks. The problems arise with CNNs when they go deeper. This is because the path for information from the input layer until the output layer (and for the gradient in the opposite direction) becomes so big, that they can getvanished before reaching the other side. DenseNets simplify the connectivity pattern between layers introduced in other architectures. DenseNets are divided into DenseBlocks, where the dimensions of the feature maps remains constant within a block, but the numberof filters between them is changed. The layers between the blocks are called Transition Layers which reduce the the number of channels to half of that of the existing channels.

Basic convolution layer with 64 filters of size 7X7 and a stride of Basic pooling layer with 3x3 max pooling and a stride of 2 .Dense Block 1 with 2 convolutions repeated 6 times.Transition layer 1 (1 Conv + 1 AvgPool). Dense Block 2 with

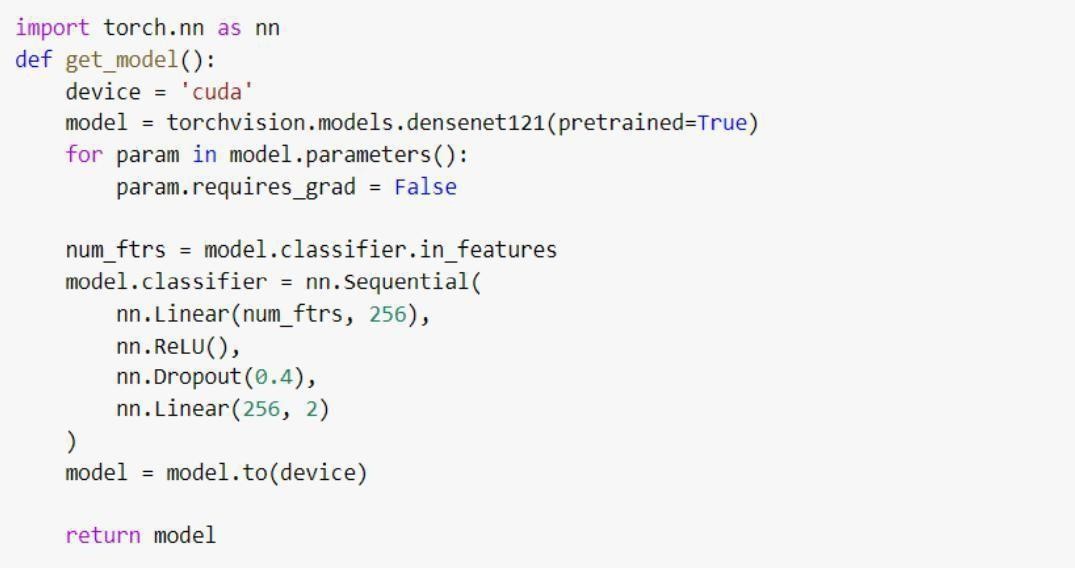
2 convolutions repeated 12 times. Transition layer 2 (1 Conv + 1 AvgPool). Dense Block 3

with 2 convolutions repeated 24 times. Transition layer 3 (1 Conv + 1 AvgPool). Dense Block 4 with 2 convolutions repeated 16 times. Global Average Pooling layer- accepts allthe feature maps of the network to perform classification output layer.



**FIG-5: Classification Layer**

* + 1. **CODE:**



**OUTPUT**:

STARTED EPOCH 27

100%|■■■■■■■■■■| 668/668 [01:36<00:00, 6.94it/s] Loss is 0.3806344535268709 Accuracy is 0.8238367016265492

100%|■■■■■■■■■■| 167/167 [00:23<00:00, 7.25it/s]

Val Loss is 0.5137113411269502 Val Accuracy is 0.8051272452234508 STARTED EPOCH 28

100%|■■■■■■■■■■| 668/668 [01:36<00:00, 6.92it/s] Loss is 0.38003217689916047 Accuracy is 0.82294785432116

100%|■■■■■■■■■■| 167/167 [00:23<00:00, 7.22it/s]

Val Loss is 0.49835072020570675 Val Accuracy is 0.8109805387651136 STARTED EPOCH 29

100%|■■■■■■■■■■| 668/668 [01:36<00:00, 6.93it/s]

Loss is 0.38024804779066296 Accuracy is 0.8239302645008008

100%|■■■■■■■■■■| 167/167 [00:23<00:00, 7.22it/s]

Val Loss is 0.5089202983650619 Val Accuracy is 0.8095883230963153 STARTED EPOCH 30

100%|■■■■■■■■■■| 668/668 [01:36<00:00, 6.95it/s]

Loss is 0.37804545350595864 Accuracy is 0.8249438622754491

100%|■■■■■■■■■■| 167/167 [00:23<00:00, 7.24it/s]

Val Loss is 0.5098220889796754 Val Accuracy is 0.8117290417591255

# CHAPTER 4 PROPOSED SYSTEM

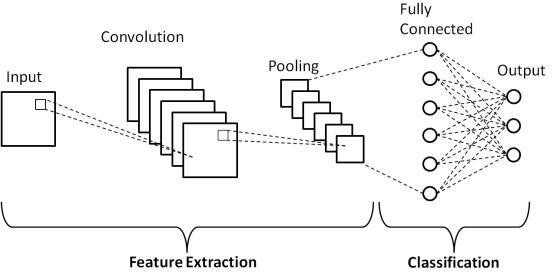
##### CNN Segmentation + connected components:

It was proposed to deal with the problems faced by the object recognition models.

Fast R-CNN is one of the state-of-the-art models at that time but it has its own challenges such as this network cannot be used in real-time.

Because it takes 2-3 seconds to predicts an image and therefore cannot be used in real-time.

The CNN segmentation and connected components will overcome the issues with R-CNN model to detect theimages.



**FIG 6: CNN segmentation architecture**

This architecture takes an image as input and resizes it to 256\*256 by keeping the aspect ratio same and performing padding.

This image is then passed in the CNN network. This model has 24 convolution layers, 4 max-pooling layers followed by 2 fully connected layers. For the reduction of the number of layers (Channels), we use 1\*1 convolution that is followed by 3\*3 convolution. Notice that the last layer of pooling predicts a cuboidal output.This is done by generating (1, 1470) from final fully connected layer and reshaping it to size (7, 7, 30).

##### Approach:

Firstly a convolutional neural network is used to segment the image, using the bounding boxes directly as amask.

Secondly connected components is used to separate multiple areas of predicted pneumonia.Finally a bounding box is simply drawn around every connected component.

##### Network:

* The network consists of a number of residual blocks with convolutions and downsampling blocks with max pooling.
* At the end of the network a single upsampling layer converts the output to the same shape as the input.

As the input to the network is 256 by 256 (instead of the original 1024 by 1024) and the network downsamplesa number of times without any meaningful upsampling (the final upsampling is just to match in 256 by 256 mask) the final prediction is very crude. If the network downsamples 4 times the final bounding boxes can onlychange with at least 16 pixels.

##### Load pneumonia locations:

Table contains [filename : pneumonia location] pairs per row.

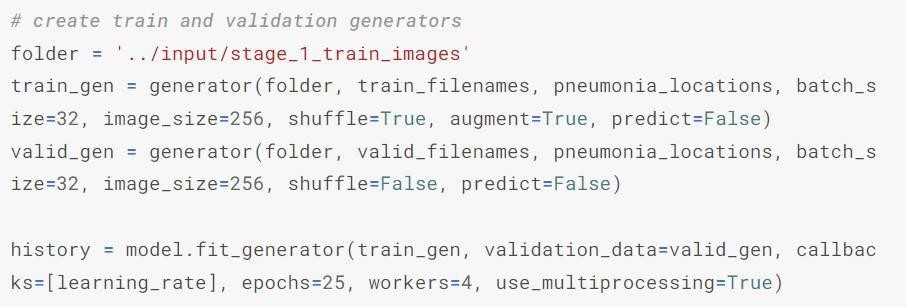
If a filename contains multiple pneumonia, the table contains multiple rows with the same filename but different pneumonia locations.

If a filename contains no pneumonia it contains a single row with an empty pneumonia location.

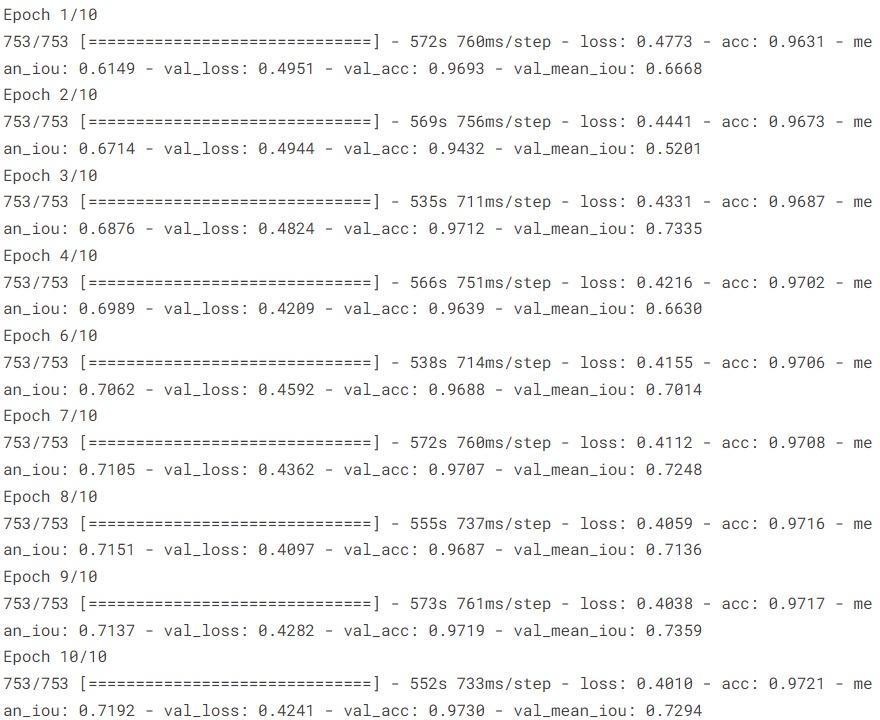
We need to load the table and transforms it into a dictionary.

* + The dictionary uses the filename as key and a list of pneumonia locations in that filename as value.
  + If a filename is not present in the dictionary it means that it contains no pneumonia.

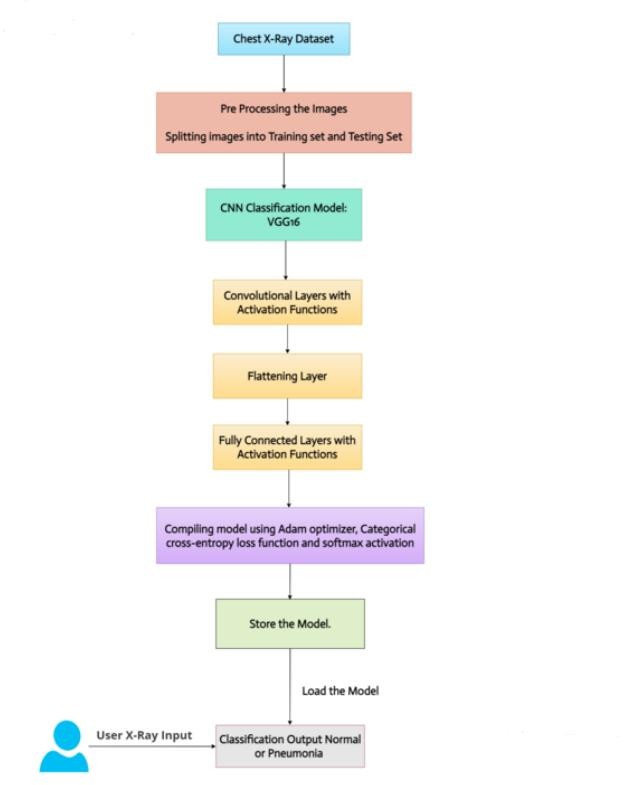
**CODE:**



**OUTPUT:**

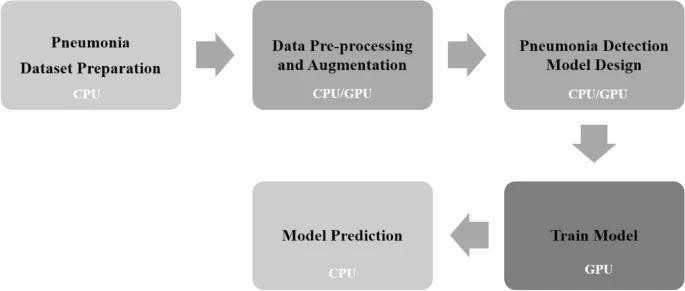


# CHAPTER 5 DESGINING



**FIG 7: Process of CNN segmentation**

The complete processes of the proposed model for Pneumonia infection identification are described further indetail. The entire method is separated into a number of stages in the following subsections, starting with collecting the images for the classification process.The proposed Pneumonia detection model consists of five phases. The complete flow of the proposed system is shown in Fig. 8.



##### FIG 8:flow chart of cnn segmentation

The complete development process of the Pneumonia infection detection system was given on subsequent subsections. It begins with dataset preparation and the preprocessing phase.

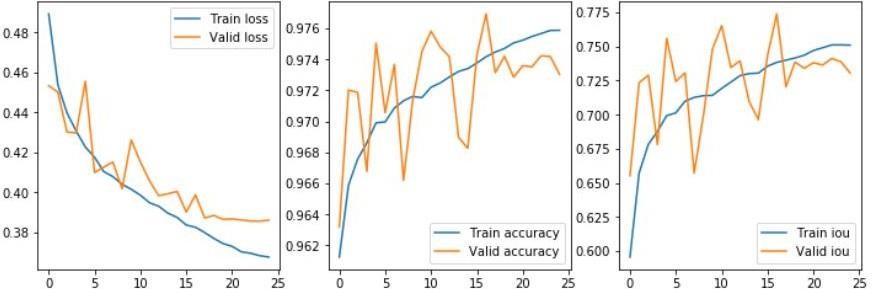
# CHAPTER 6 RESULTS AND DISCUSSION

#### RESULT:

* **IOU LOSS:**

Intersection-Over-Union is **a common evaluation metric for semantic image segmentation**. For an individual class, the IoU metric is defined as follows: iou = true\_positives / (true\_positives + false\_positives + false\_negatives)

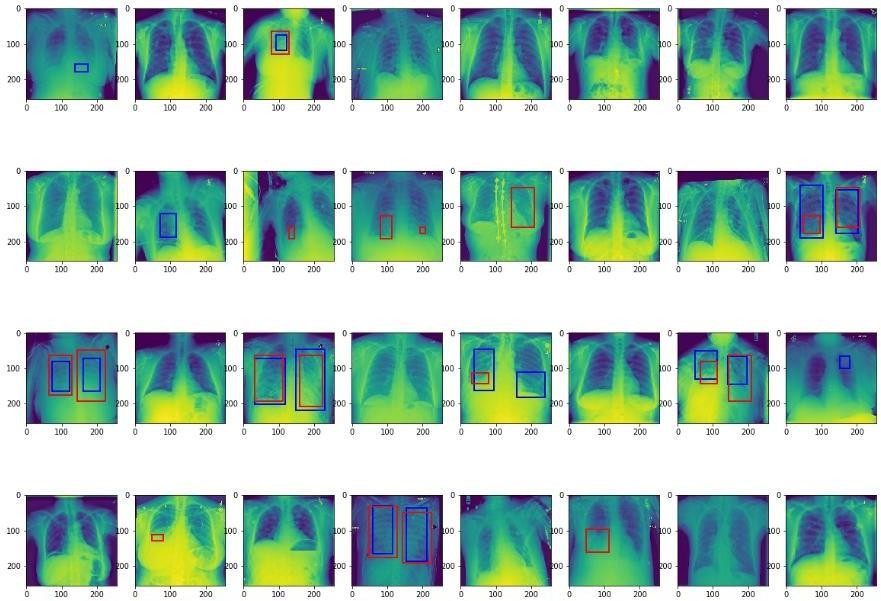
IoU threshold : Intersection over Union, a value used in object detection to measure the overlap of a predictedversus actual bounding box for an object. The closer the predicted bounding box values are to the actual bounding box values the greater the intersection, and the greater the IoU value.



**FIG 9: graphs for accuracy**

We got training accuracy as 97.21 % with train loss as 0.4010 and training mean iou as 0.7192

We got validation accuracy as 97.30 % with validation loss as 0.4241 and validation mean iou as 0.7294



**FIG 10: Prediction of Bounding Box for a batch**

In this image we have classified pneumonia and non pneumonia images the images which contain both valid and train loss in high the images are pneumonia and images with less loss is non pneumonia

From the above image first row third image is pneumonia ,secound row secound and fifth images are pneumonia images

Remaining images with less loss of validity are non pneumoina images

#### DECISION:

The customized model i.e a combination of CNN based feature-extraction and supervised classifier algorithm resulted in optimal solution for classifying abnormal (Pneumonia labeled) and normal Chest X-Ray images primarily due to the substantive features provided by DenseNets followed by optimal hyper-parameter values of SVM classifier. Literature studies reveal the contribution of transfer learning methods including feature- extractions toward visual recognition tasks For this reason, we extracted features from various variants of pre- trained CNN models available such as VGGNets Inception ResNet- 50 and DenseNets . Studies from the literature also reveal the use of classifiers in combination with CNN- based feature extraction majorly in medicalimage analysis to meliorate the performance of models.

Following the mentioned past approaches, we evaluated each of the pre-trained models with distinct classifiers to determine the ideal model for the purpose. We observed from the comparative experimental results presentedin and that ResNet50 outperformed the results of all the other pre-trained CNN models when employed with default parameter values of SVM classifier. In addition, DenseNets were also observed to achieve results close to ResNet50. Literature studies reveal that DenseNets outperformed all the pre-trained CNNs in the ImageNet dataset . For this reason, we chose ResNet50, DenseNet-121 and DenseNet169 as the optimal CNN models for the feature-extraction stage and SVM as the optimal classifier for the classification stage for further experimentsin the study. The selection of SVM classifier with rbf kernel based on the statistical results presented in further led to hunt of optimal hyperparameter values (C and gamma). In the process of tuning hyper-parameters, we performed close to 350 combinations of C and gamma, the crucial combinations among these are presented in. We observed in this process that DenseNet-169 outperformed all the other customized models and hence chosenas the best feature-extractor for the final customized model followed by optimal hyper-parameter values of SVM rbf kernel.

The best results achieved with DenseNet169 architecture as feature extractors can be explained due to its capability of accessing feature-maps from all of its preceding layers. Literature studies of DenseNets mentions the information flow from the beginning layer to the end layers and removal of redundant features by transition layers as the primary reasons for the high-features representations. To our knowledge, no literature was found toperform the studies on the combination of CNN based feature extractions and supervised classifier algorithms for the underlying task. In regard, we have proposed a model architecture for detecting Pneumonia from frontal chest X-ray images with the utilization of Densenet as feature-extractors and SVM as the process of melioratingthe model performance, we found that our customized model outperforms the results documented in the recentlyreleased work of Benjamin for the same problem of pneumonia detection.

# CHAPTER 7 CONCLUSION AND FUTURE

**WORK**

##### Future Work:

Future Enhancement is being planned to further analyze and enhance the protocol is to a bit faster,but it reduces the credibility of the whole system by being partially centralized because it only runs where the authority wants it.

#### CONCLUSION :

We develop an algorithm which detects pneumonia from frontal-view chest X-ray images at a level exceeding practicing radiologist. We also show that an easy extension of our algorithm to detect multiple diseases outperforms previous state of the art on ChestX-ray14, the most important publicly available chest X-ray dataset. With automation at the extent of experts, we hope that this technology can improve health care delivery and increase access to medical imaging expertise in parts of the planet where access to skilled radiologists is limited. An efficient model is made using deep learning algorithms in Python language which can help doctors to detect this deadly disease. The proposed framework is compared with state-of-the art methods of machine learning and found to be more efficient in prediction with an average accuracy of 84%, which is found to be better than all other

classifiers.

# CHAPTER 8 REFERENCES

1. Rahman T., Chowdhury M.E., Khandakar A., Islam K.R., Islam K.F., Mahbub Z.B., Kadir M.A., Kashem S. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray.
2. Nicholas E Ross, Charles J Pritchard, David M Rubin and Adriano G Duse, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears", Medical and Biological Engineering and Computing, vol. 44, no. 5, 2006.
3. Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition", 2014.
4. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7345724/>
5. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0256630
6. [https://www.hindawi.com/journals/wcmc/2021/5556635/#materials-and-methods](https://www.hindawi.com/journals/wcmc/2021/5556635/)
7. https://ieeexplore.ieee.org/document/8869364